

A strengthened entropy power inequality for log-concave densities

Giuseppe Toscani *

August 19, 2014

Abstract. We show that Shannon's entropy-power inequality admits a strengthened version in the case in which the densities are log-concave. In such a case, in fact, one can extend the Blachman–Stam argument [25, 5] to obtain a sharp inequality for the second derivative of Shannon's entropy functional with respect to the heat semigroup.

Keywords. Entropy-power inequality, Blachman–Stam inequality.

1 Introduction

Given a random vector X in \mathbb{R}^n , $n \geq 1$, with a smooth, rapidly decaying probability density $f(x)$ such that $\log f$ has growth at most polynomial at infinity, consider the functional

$$\begin{aligned} J(X) = J(f) &= \sum_{i,j=1}^n \int_{\{f>0\}} [\partial_{ij}(\log f)]^2 f \, dx = \\ &= \sum_{i,j=1}^n \int_{\{f>0\}} \left[\frac{\partial_{ij} f}{f} - \frac{\partial_i f \partial_j f}{f^2} \right]^2 f \, dx. \end{aligned} \quad (1)$$

This functional is closely related to the entropy functional (or Shannon's entropy) of the random vector X

$$H(X) = H(f) = - \int_{\mathbb{R}^n} f(x) \log f(x) \, dx. \quad (2)$$

*Department of Mathematics, University of Pavia, via Ferrata 1, 27100 Pavia, Italy.
giuseppe.toscani@unipv.it

through the solution $u(x, t)$ of the initial value problem for the heat equation in the whole space \mathbb{R}^n ,

$$\frac{\partial u}{\partial t} = \kappa \Delta u, \quad u(x, t = 0) = f(x), \quad (3)$$

where κ is the diffusion constant. To our knowledge, in one-dimension this link was first remarked in 1965 by McKean [19], who computed the evolution in time of the subsequent derivatives of the entropy functional $H(u(t))$ of the solution to (3), with the diffusion constant $\kappa = 1$. McKean computations at the first two orders gave the identities

$$I(f) = \left. \frac{d}{dt} \right|_{t=0} H(u(t)); \quad J(f) = -\frac{1}{2} \left. \frac{d}{dt} \right|_{t=0} I(u(t)). \quad (4)$$

In (4) $I(f)$ defines, for a given random vector X in \mathbb{R}^n , its Fisher information

$$I(X) = I(f) = \int_{\{f>0\}} \frac{|\nabla f(x)|^2}{f(x)} dx. \quad (5)$$

Hence, $J(u(t))$ can be interpreted as a measure of the (nonnegative) *entropy production* of the Fisher information when evaluated along the solution to the heat equation.

While the importance of the first identity in (4), well-known in information theory with the name of DeBruijn's identity is well established [9], the role of the second identity in (4) seems to be restricted to its use in the proof of the concavity of entropy power [8], and to the so-called entropy-entropy methods [1, 3], where it has been shown highly useful in the proof of the logarithmic Sobolev inequality [26].

In this paper we study inequalities related to the functional $J(\cdot)$, when evaluated on convolutions. The main result is a new inequality for $J(X + Y)$, where X and Y are independent random vectors in \mathbb{R}^n , such that their probability densities f and g are *log-concave*, and $J(X)$, $J(Y)$ are well defined. By resorting to an argument close to that used by Blachman [5] in his original proof of entropy power inequality, for any given pair of positive constants a, b , we prove the bound

$$J(f * g) \leq \frac{a^4}{(a+b)^4} J(f) + \frac{b^4}{(a+b)^4} J(g) + \frac{2a^2b^2}{(a+b)^4} H(f, g), \quad (6)$$

where

$$H(f, g) = \sum_{i,j=1}^n \int_{\{f>0\}} \frac{\partial_i f \partial_j f}{f} dx \int_{\{g>0\}} \frac{\partial_i g \partial_j g}{g} dx. \quad (7)$$

Note that, in one-dimension $H(f, g)$ coincides with the product of the Fisher information of f and g , $H(f, g) = I(f)I(g)$. Inequality (6) is sharp. Indeed, there is equality if and only if X and Y are n -dimensional Gaussian vectors with covariance matrices proportional to aI_n and bI_n respectively, where I_n is the identity matrix.

Even if inequality (6) is restricted to the set of log-concave densities, this set includes many of the most commonly-encountered parametric families of probability density functions [20]. Among other properties, log-concave densities are stable under convolution. If f and g are (possibly multidimensional) log-concave densities, then their convolution $f * g$ is log-concave.

Optimizing over a and b , one obtains from (6) the inequality

$$\frac{1}{\sqrt{J(X+Y)}} \geq \frac{1}{\sqrt{J(X)}} + \frac{1}{\sqrt{J(Y)}}, \quad (8)$$

where, also in this case, equality holds if and only if both X and Y are Gaussian random vectors with proportional covariance matrices.

Inequality (8) shows that, at least if applied to log-concave probability densities, the functional $J(\cdot)$ behaves with respect to convolutions like Shannon's entropy power [24, 25] and Fisher information [25, 5]. Actually, Shannon's entropy power inequality, due to Shannon and Stam [24, 25] (cf. also [8, 15, 16, 22, 31] for other proofs and extensions) gives a lower bound on Shannon's entropy power of the sum of independent random variables X, Y with values in \mathbb{R}^n with densities

$$N(X+Y) \geq N(X) + N(Y), \quad (9)$$

with equality if and only if X and Y are Gaussian random vectors with proportional covariance matrices. In (9) the entropy-power of the random variable X with values in \mathbb{R}^n is defined by

$$N(X) = N(f) = \exp\left(\frac{2}{n}H(X)\right). \quad (10)$$

With no doubts, this is one of the fundamental information theoretic inequalities [13].

Likewise, Blachman–Stam inequality [25, 5] gives a lower bound on the inverse of Fisher information of the sum of independent random vectors with (smooth) densities

$$\frac{1}{I(X+Y)} \geq \frac{1}{I(X)} + \frac{1}{I(Y)}, \quad (11)$$

still with equality if and only if X and Y are Gaussian random vectors with proportional covariance matrices.

The fact that inequalities (8), (9) and (11) share a common nature is clarified by noticing that, when evaluated in correspondence to a Gaussian vector X with covariance matrix σI_n , the three (related by (4)) functionals $N(X)$, $1/I(X)$ and $1/\sqrt{J(X)}$ are linear functions of σ .

An interesting application of inequality (6) is linked to the evolution in time of the functional

$$\Lambda(t) = \Lambda(f(t), g(t)) = H(f(t) * g(t)) - \kappa H(f(t)) - (1 - \kappa)H(g(t)), \quad (12)$$

where, for a positive constant κ , with $0 < \kappa < 1$, $f(x, t)$ (respectively $g(x, t)$) are the solutions to the heat equation (3) with diffusion constant κ (respectively $1 - \kappa$), corresponding to the initial data f and g which are log-concave probability densities in \mathbb{R}^n . Since, for $\alpha > 0$

$$H(f_\alpha) = H(f) - \log \alpha,$$

the functional $\Lambda(t)$ is dilation invariant, that is invariant with respect to the scaling

$$f(x) \rightarrow f_\alpha(x) = \alpha^n f(\alpha x), \quad \alpha > 0. \quad (13)$$

In [29] we proved that $\Lambda(t)$ is monotonically decreasing in time from $\Lambda(0)$ to

$$\lim_{t \rightarrow \infty} \Lambda(t) = -\frac{n}{2} [\kappa \log \kappa + (1 - \kappa) \log(1 - \kappa)],$$

which implies the inequality

$$H(f * g) - \kappa H(f) - (1 - \kappa) H(g) \geq -\frac{n}{2} [\kappa \log \kappa + (1 - \kappa) \log(1 - \kappa)].$$

By optimizing over κ , this inequality implies the entropy power inequality (9). By choosing now log-concave densities f and g as initial data in the heat equations, we can prove, in consequence of inequality (6), that the functional $\Lambda(t)$ is a convex function of time, and this implies, optimizing over κ , the strengthened entropy power inequality

$$N(X + Y) \geq [N(X) + N(Y)] R(X, Y), \quad (14)$$

where the quantity $R(X, Y) \geq 1$ can be interpreted as a measure of the *non-Gaussianity* of the two random vectors X, Y . Indeed, $R(X, Y) = 1$ if and only if both X and Y are Gaussian random vectors.

Inequality (12) describes a new property of convolutions of Gaussian densities. Among all solutions to the heat equation with log-concave densities as initial data, the Gaussian self-similar solutions are the unique ones for which the functional $\Lambda(t)$ remains constant in time. In all the other cases $\Lambda(t)$ is a *convex* function of time. This property is reminiscent of the well-known *concavity of entropy power* theorem, which asserts that, given the solution $u(x, t)$ to the heat equation (3),

$$\frac{d^2}{dt^2} (N(u(t))) \leq 0. \quad (15)$$

Inequality (15) is due to Costa [8]. Later, the proof has been simplified in [12, 13], by an argument based on the Blachman–Stam inequality [5]. Moreover, a short and simple proof has been obtained by Villani [30], using ideas from the aforementioned McKean paper [19]. It is interesting to notice that the concavity property has been extended to Renyi entropy power, when evaluated along the solution to a nonlinear diffusion equation [23].

These studies reveal that the heat equation (and, more in general, the nonlinear diffusion equations) are quite useful instruments to be used in connection with the proof of inequalities. In some recent papers [27, 28, 29], this connection has been made evident by showing that, in addition to Shannon's entropy power inequality, a number of inequalities in sharp form, like the classical Young's inequality and its converse [6], the Brascamp–Lieb type inequalities [6], Babenko's inequality [2], Prékopa–Leindler inequality [21], Nash's inequality and the logarithmic Sobolev inequality follow by monotonicity arguments of various Lyapunov functionals, when evaluated on solutions to the heat equation.

A careful reading of Dembo's proof of the concavity of entropy power [12] clarifies once more the connections among the functionals H, I and J . Actually, Dembo proved inequality (15) in the equivalent form

$$J(f) \geq \frac{1}{n} I(f)^2, \quad J(f) = -\frac{1}{2} \left. \frac{d}{dt} \right|_{t=0} I(f * M_{2t}), \quad (16)$$

that is reminiscent of (4).

As already mentioned in this introduction, while the functional $J(f)$ has been introduced in Shannon's theory in connection with the proof of the concavity of Shannon's entropy power, so that it appears in the related literature after Costa's paper [8] in 1985, in one dimension of space, various properties of the $J(f)$ functional were considered by McKean in its pioneering paper on Kac's caricature of Maxwell molecules in kinetic theory of rarefied gases [19] in 1965. In his paper, McKean investigated various connections between Fisher information (5) and its derivative $J(f)$ along the heat flux, motivated by proving the old conjecture that subsequent derivatives of Shannon's entropy along the solution to the heat equation alternate in sign. McKean original inequalities for the functional $J(f)$ were subsequently generalized to higher dimensions to give a new proof of the logarithmic Sobolev inequality with an explicit remainder [26].

In more details, Section 2 will be devoted to the proof of inequality (6). For the sake of clarity, we will present first the proof in one dimension of space. Then, a multi-dimensional version will be obtained resorting to some additional result. In Section 3 we will show how inequality (6) could be fruitfully used to obtain, for log-concave densities, the strengthened version (14) of the entropy power inequality. Unfortunately, it seems quite difficult to express the additional term $R(X, Y)$ in (14) in the form of a *distance* of the involved densities from the Gaussian density, and we live we leave this question to future research.

2 A new inequality for convolutions

2.1 Log-concave functions and scores

We recall that a function f on \mathbb{R}^n is log-concave if it is of the form

$$f(x) = \exp \{ -\Phi(x) \}, \quad (17)$$

for some convex function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$. A prime example is the Gaussian density, where $\Phi(x)$ is quadratic in x . Further, log-concave distributions include Gamma distributions with shape parameter at least one, $Beta(\alpha, \beta)$ distributions with $\alpha, \beta \geq 1$, Weibull distributions with shape parameter at least one, Gumbel, logistic and Laplace densities (see, for example, Marshall and Olkin [20]). Log-concave functions have a number of properties that are desirable for modelling. Marginal distributions, convolutions and product measures of log-concave distributions and densities are again log-concave (cf. for example, Dharmadhikari and Joag-Dev [11]).

A main consequence of log-concavity, which is at the basis of most computations in this paper, is the following. Consider the heat equation (3) in \mathbb{R}^n , $n \geq 1$. If $M_\sigma(x)$ denotes the Gaussian density in \mathbb{R}^n with zero mean and covariance matrix σI_n

$$M_\sigma(x) = \frac{1}{(2\pi\sigma)^{n/2}} \exp \left(-\frac{|x|^2}{2\sigma} \right), \quad (18)$$

the solution at time t to the heat equation (3) coincides with $u = f * M_{2\kappa t}$.

Assume that the initial datum $f(x)$ is a non-negative, log-concave integrable function. Then, at each subsequent time $t > 0$, the solution $u(\cdot, t)$ to the heat equation, convolution of the log-concave functions f and the Gaussian density $M_{2\kappa t}$ defined in (18), is a non-negative integrable log-concave function. In other words, the heat equation propagates log-concavity.

This simple remark, allows to proof things by using smooth functions with fast decay at infinity.

It is interesting to notice that the expressions of Shannon's entropy H , Fisher information I and Fisher's entropy production J take a very simple form if evaluated in correspondence to log-concave densities f , when written as in (17). In this case, if X is a random vector in \mathbb{R}^n with density f , these functionals can be easily recognized as moments of $\Phi(X)$ or of its derivatives. It is immediate to reckon that Shannon's entropy H coincides with

$$H(f) = \int_{\mathbb{R}^n} \Phi(x) f(x) dx. \quad (19)$$

The Fisher information I reads

$$I(f) = \int_{\mathbb{R}^n} |\nabla \Phi(x)|^2 f(x) dx = \sum_{i=1}^n \int_{\mathbb{R}^n} |\partial_i \Phi(x)|^2 f(x) dx, \quad (20)$$

and, last, Fisher's entropy production J takes the form

$$J(f) = \sum_{i,j=1}^n \int_{\mathbb{R}^n} |\partial_{ij}\Phi(x)|^2 f(x) dx. \quad (21)$$

Thus, the functionals are well-defined in terms of the convex function Φ characterizing the log-concave function f .

For the log-concave Gaussian density (18)

$$\Phi(x) = \frac{|x|^2}{2\sigma} + \frac{n}{2} \log 2\pi\sigma,$$

which implies, for $i, j = 1, 2, \dots, n$

$$\partial_i \Phi(x) = \frac{x_i}{\sigma}, \quad \partial_{ij} \Phi(x) = \frac{1}{\sigma} \delta_{ij},$$

where, as usual, δ_{ij} is the Kronecker delta.

According to the standard definition, given a random vector X in \mathbb{R}^n distributed with with absolutely continuous probability density function $f(x)$

$$\rho(X) = \frac{\nabla f(X)}{f(X)}, \quad (22)$$

denotes the (almost everywhere defined) score function of the random variable [10] (cf. also [18] for further details). The score has zero mean, and its variance is just the Fisher information. For log-concave densities, which are expressed in the form (17)

$$\rho(X) = -\nabla \Phi(X) \quad (23)$$

In view of definition (1) and (22) one can think to introduce the concept of second-order score of a random vector X in \mathbb{R}^n , defined by the symmetric Hessian matrix $\mathcal{H}(X)$ of $-\log f(X)$, with elements

$$\Psi_{ij}(X) = \frac{\partial_i f \partial_j f(X)}{f^2(X)} - \frac{\partial_{ij} f(X)}{f(X)}. \quad (24)$$

Then, as the Fisher information coincides the second moment of the score function, the functional $J(X)$ in (1) is expressed by the moment of the trace of the product matrix $\mathcal{H}(X) \cdot \mathcal{H}(X)$. For a log-concave function, the element $\Psi_{i,j}$ of the Hessian matrix $\mathcal{H}(X)$ defining the second-order score function takes the simple expression

$$\Psi_{ij}(X) = \partial_{ij} \Phi(X). \quad (25)$$

Note that a Gaussian vector M_σ is uniquely defined by a *linear* score function $\rho(M_\sigma) = M_\sigma/\sigma$ and by a *constant* second-order score matrix $\mathcal{H}(M_\sigma) = I_n/\sigma$.

2.2 The one-dimensional case

For the moment, let us fix $n = 1$. In the rest of this section, we will only consider smooth log-concave probability densities $f(x)$ (cf. definition (17)) such that $\Phi(x) = -\log f(x)$ has growth at most polynomial at infinity. In order not to worry about derivatives of logarithms, which will often appear in the proof, we may also impose that $|\Phi'(x)/\Phi(x)| \leq C(1 + |x|^2)$ for some positive constant C . The general case will easily follow by a density argument [17]. Let

$$k(x) = f * g(x).$$

The main argument here is due to Blachman [5], who proved in this way inequality (11). Since for any pair of positive constants a, b we have the identity

$$(a + b)k'(x) = a \int_{\mathbb{R}} f'(x - y)g(y) dy + b \int_{\mathbb{R}} f(x - y)g'(y) dy,$$

dividing by $k(x) > 0$ we obtain

$$(a + b) \frac{k'(x)}{k(x)} = a \int_{\mathbb{R}} \frac{f'(x - y)}{f(x - y)} \frac{f(x - y)g(y)}{k(x)} dy + b \int_{\mathbb{R}} \frac{g'(y)}{g(y)} \frac{f(x - y)g(y)}{k(x)} dy =$$

$$\int_{\mathbb{R}} \left(a \frac{f'(x - y)}{f(x - y)} + b \frac{g'(y)}{g(y)} \right) d\mu_x(y).$$

We denoted

$$d\mu_x(y) = \frac{f(x - y)g(y)}{k(x)} dy.$$

Note that, for every $x \in \mathbb{R}$, $d\mu_x$ is a unit measure on \mathbb{R} . Consequently, by Jensen's inequality

$$(a + b)^2 \left(\frac{k'(x)}{k(x)} \right)^2 = \left[\int_{\mathbb{R}} \left(a \frac{f'(x - y)}{f(x - y)} + b \frac{g'(y)}{g(y)} \right) d\mu_x(y) \right]^2 \leq$$

$$\int_{\mathbb{R}} \left(a \frac{f'(x - y)}{f(x - y)} + b \frac{g'(y)}{g(y)} \right)^2 d\mu_x(y). \quad (26)$$

On the other hand, by analogous argument, for any pair of positive constants a, b we have the identity

$$(a + b)^2 k''(x) = a^2 \int_{\mathbb{R}} f''(x - y)g(y) dy + b^2 \int_{\mathbb{R}} f(x - y)g''(y) dy + 2ab \int_{\mathbb{R}} f'(x - y)g'(y) dy.$$

Thus, dividing again by $k(x) > 0$ we obtain

$$(a + b)^2 \frac{k''(x)}{k(x)} = \int_{\mathbb{R}} \left(a^2 \frac{f''(x - y)}{f(x - y)} + b^2 \frac{g''(y)}{g(y)} + 2ab \frac{f'(x - y)}{f(x - y)} \frac{g'(y)}{g(y)} \right) d\mu_x(y). \quad (27)$$

If we subtract identity (27) from inequality (26) we conclude with the inequality

$$(a+b)^2 \left[\left(\frac{k'(x)}{k(x)} \right)^2 - \frac{k''(x)}{k(x)} \right] \leq \int_{\mathbb{R}} \left\{ a^2 \left[\left(\frac{f'(x-y)}{f(x-y)} \right)^2 - \frac{f''(x-y)}{f(x-y)} \right] + b^2 \left[\left(\frac{g'(y)}{g(y)} \right)^2 - \frac{g''(y)}{g(y)} \right] \right\} d\mu_x(y). \quad (28)$$

It is important to note that, since the functions f, g (and consequently k) are log-concave, the left-hand side of inequality (28) is non-negative. Therefore, taking the square on both sides of (28), and using once more Jensen's inequality we end up with the inequality

$$(a+b)^4 \left[\left(\frac{k'(x)}{k(x)} \right)^2 - \frac{k''(x)}{k(x)} \right]^2 \leq \int_{\mathbb{R}} \left\{ a^2 \left[\left(\frac{f'(x-y)}{f(x-y)} \right)^2 - \frac{f''(x-y)}{f(x-y)} \right] + b^2 \left[\left(\frac{g'(y)}{g(y)} \right)^2 - \frac{g''(y)}{g(y)} \right] \right\}^2 d\mu_x(y). \quad (29)$$

Multiplying both sides of (29) by $k(x)$, and integrating over \mathbb{R} yields the inequality

$$(a+b)^4 J(k) \leq a^4 J(f) + b^4 J(g) + 2a^2 b^2 I(f)I(g).$$

Indeed, in one dimension, definition (1) of the functional $J(\cdot)$ reduces to

$$J(f) = \int_{\{f>0\}} \left[\left(\frac{f'(x)}{f(x)} \right)^2 - \frac{f''(x)}{f(x)} \right]^2 f(x) dx. \quad (30)$$

Moreover

$$\begin{aligned} \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy \left[\left(\frac{f'(x-y)}{f(x-y)} \right)^2 - \frac{f''(x-y)}{f(x-y)} \right] \left[\left(\frac{g'(y)}{g(y)} \right)^2 - \frac{g''(y)}{g(y)} \right] f(x-y)g(y) = \\ \int_{\mathbb{R}} dx \int_{\mathbb{R}} dy \left(\frac{f'(x-y)}{f(x-y)} \right)^2 \left(\frac{g'(y)}{g(y)} \right)^2 f(x-y)g(y) = I(f)I(g), \end{aligned}$$

where $I(f)$ (respectively $I(g)$) denotes the Fisher information of f (respectively g)

$$I(f) = \int_{\{f>0\}} \frac{(f'(x))^2}{f(x)} dx.$$

The cases of equality in (26) and (29) are easily found resorting to the following argument. Equality follows if, after application of Jensen's inequality, there is equality

in (26). On the other hand, for any convex function φ and unit measure $d\mu$ on the set Ω , equality in Jensen's inequality

$$\varphi\left(\int_{\Omega} f d\mu\right) \leq \int_{\Omega} \varphi(f) d\mu$$

holds true if and only if f is constant, so that

$$f = \int_{\Omega} f d\mu.$$

In our case, this means that there is equality if and only if the function

$$a \frac{f'(x-y)}{f(x-y)} + b \frac{g'(y)}{g(y)}$$

does not depend on y . If this is the case, taking the derivative with respect to y , and using the identity

$$\frac{d}{dy} \left(\frac{f'(x-y)}{f(x-y)} \right) = -\frac{d}{dx} \left(\frac{f'(x-y)}{f(x-y)} \right),$$

we conclude that f and g are forced to satisfy

$$a \frac{d^2}{dx^2} \log f(x-y) = b \frac{d^2}{dy^2} \log g(y). \quad (31)$$

Note that (31) can be verified if and only if the functions on both sides are constant. Thus, there is equality if and only if

$$\log f(x) = b_1 x^2 + c_1 x + d_1, \quad \log g(x) = b_2 x^2 + c_2 x + d_2. \quad (32)$$

By coupling (32) with (31), we obtain that there is equality in (5) if and only if f and g are Gaussian densities, of variances ca and cb , respectively, for any given positive constant c . Analogous argument leads to the same conclusion as far as inequality (29) is concerned. We proved

Theorem 1. *Let $f(x)$ and $g(x)$ be log-concave probability density functions with values in \mathbb{R} , such that both $J(f)$ and $J(g)$, as given by (30) are bounded. Then, also $J(f * g)$ is bounded, and for any pair of positive constants a, b*

$$J(f * g) \leq \frac{a^4}{(a+b)^4} J(f) + \frac{b^4}{(a+b)^4} J(g) + 2 \frac{a^2 b^2}{(a+b)^4} I(f) I(g). \quad (33)$$

Moreover, there is equality in (33) if and only if, up to translation and dilation f and g are Gaussian densities, $f(x) = M_a(x)$ and $g(x) = M_b(x)$.

Remark 2. The condition of log-concavity enters into the proof of Lemma 1 when we pass from inequality (28) to inequality (29). Without the condition of log-concavity, in fact, the left-hand side of (28) has no sign, and (29) does not hold true. Of course, this fact does not exclude the possibility that inequality (33) could hold also for other classes of probability densities, but if any, another method of proof has to be found, or a counterexample is needed.

Theorem 1 allows to prove inequality (8). To this aim, note that, for any pair of positive constants a, b

$$2\sqrt{J(f)J(g)} \leq \frac{a}{b}J(f) + \frac{b}{a}J(g).$$

Moreover, as proven first by Dembo [12], and later on by Villani [30] with a proof based on McKean ideas [19] (16) implies

$$J(f) \geq I(f)^2. \quad (34)$$

Remark 3. The proof of (34) is immediate and enlightening. Given the random variable X distributed with a sufficiently smooth density $f(x)$, consider the (almost everywhere defined) second-order score variable (cf. definition (24))

$$\Psi(X) = \left(\frac{f'(X)}{f(X)} \right)^2 - \frac{f''(X)}{f(X)}. \quad (35)$$

Then, denoting with $\langle Y \rangle$ the mathematical expectation of the random variable Y , it holds

$$I(f) = I(X) = \langle \Psi(X) \rangle, \quad J(f) = J(X) = \langle \Psi(X)^2 \rangle.$$

Then, (34) coincides with the standard inequality $\langle \Psi(X)^2 \rangle \geq \langle \Psi(X) \rangle^2$. Note moreover that equality in (34) holds if and only if $\Psi(X)$ is constant, or, what is the same, if

$$\frac{d^2}{dx^2} \log f(x) = c.$$

As observed in the proof of Theorem 1 this implies that X is a Gaussian variable.

Grace to inequality (34)

$$2I(f)I(g) \leq 2\sqrt{J(f)J(g)} \leq \frac{a}{b}J(f) + \frac{b}{a}J(g). \quad (36)$$

Using (36) to bound from above the last term in inequality (33) we obtain

$$J(f * g) \leq \frac{a^3}{(a+b)^3}J(f) + \frac{b^3}{(a+b)^3}J(g). \quad (37)$$

Optimizing over $z = a/(a + b)$, with $z \in [0, 1]$, one finds that the minimum of the right-hand side is obtained when

$$z = \bar{z} = \frac{\sqrt{J(g)}}{\sqrt{J(f)} + \sqrt{J(g)}}, \quad (38)$$

which implies inequality (8). Thus we proved

Corollary 4. *Let X and Y be independent random variables with log-concave probability density functions with values in \mathbb{R} , such that both $J(X)$ and $J(Y)$, as given by (30) are bounded. Then*

$$\frac{1}{\sqrt{J(X+Y)}} \geq \frac{1}{\sqrt{J(X)}} + \frac{1}{\sqrt{J(Y)}}.$$

Moreover, there is equality if and only if, up to translation and dilation X and Y are Gaussian variables.

Remark 5. Inequality (33) implies in general a stronger inequality. In fact, to obtain inequality (37) we discarded the (non-positive) term

$$-R(f, g, a, b) = \frac{a^2 b^2}{(a+b)^4} \left(2I(f)I(g) - \frac{a}{b}J(f) - \frac{b}{a}J(g) \right). \quad (39)$$

By evaluating the value of $R(f, g, a, b)$ in $z = \bar{z}$, one shows that inequality (33) is improved by the following

$$\frac{1}{\sqrt{J(X+Y)}} \geq \left(\frac{1}{\sqrt{J(X)}} + \frac{1}{\sqrt{J(Y)}} \right) \mathcal{R}(X, Y), \quad (40)$$

where

$$1 \leq \mathcal{R}(X, Y) = \left(1 - 2 \frac{\sqrt{J(X)}\sqrt{J(Y)} - I(X)I(Y)}{(\sqrt{J(X)} + \sqrt{J(Y)})^2} \right)^{-1/2}.$$

As before, $\mathcal{R}(X, Y) = 1$ if and only if X and Y are Gaussian random variables.

Note that the non-negative remainder $R(f, g, a, b)$ can be bounded from below in terms of other expressions. In particular, one of these bounds is particularly significative. Adding and subtracting to the right-hand side of (39) the positive quantity $aI^2(f)/b + bI^2(g)/a$ one obtains the bound

$$R(f, g, a, b) \geq \frac{a^2 b^2}{(a+b)^4} \left[\frac{a}{b}(J(f) - I^2(f)) - \frac{b}{a}(J(g) - I^2(g)) \right]. \quad (41)$$

This implies that (37) can be improved by the following

$$\begin{aligned} J(f * g) &\leq \frac{a^3}{(a+b)^3} J(f) + \frac{b^3}{(a+b)^3} J(g) + \\ &- \frac{a^2 b^2}{(a+b)^4} \left[\frac{a}{b}(J(f) - I^2(f)) - \frac{b}{a}(J(g) - I^2(g)) \right]. \end{aligned} \quad (42)$$

2.3 The general case

With few variants, the proof in the multi-dimensional case follows along the same lines of the one-dimensional one. Let $f(x)$ and $g(x)$, with $x \in \mathbb{R}^n$ be multidimensional log-concave functions, and let $k(x) = f * g(x)$ be their log-concave convolution. In addition, let us suppose that both $f(x)$ and $g(x)$ are sufficiently smooth and decay at infinity in such a way to justify computations. To simplify notations, given a function $f(x)$, with $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, $n > 1$, we denote its partial derivatives as

$$f_i(x) = \partial_i f(x), \quad f_{ij}(x) = \partial_{ij} f(x).$$

For any given vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, and positive constants a, b we have the identity

$$(a + b) \sum_{i=1}^n \alpha_i \frac{k_i(x)}{k(x)} = \int_{\mathbb{R}^n} \sum_{i=1}^n \alpha_i \left(a \frac{f_i(x-y)}{f(x-y)} + b \frac{g_i(y)}{g(y)} \right) d\mu_x(y),$$

where now, for every $x \in \mathbb{R}^n$

$$d\mu_x(y) = \frac{f(x-y)g(y)}{k(x)} dy,$$

is a unit measure on \mathbb{R}^n . Therefore, by Jensen's inequality

$$(a + b)^2 \left(\sum_{i=1}^n \alpha_i \frac{k_i(x)}{k(x)} \right)^2 = (a + b)^2 \sum_{i,j=1}^n \alpha_i \alpha_j \frac{k_i(x)}{k(x)} \frac{k_j(x)}{k(x)} \leq \int_{\mathbb{R}^n} \sum_{i,j=1}^n \alpha_i \alpha_j \left(a^2 \frac{f_i(x-y)}{f(x-y)} \frac{f_j(x-y)}{f(x-y)} + b^2 \frac{g_i(y)}{g(y)} \frac{g_j(y)}{g(y)} + 2ab \frac{f_i(x-y)}{f(x-y)} \frac{g_j(y)}{g(y)} \right) d\mu_x(y).$$

Likewise, thanks to the identity

$$(a + b)^2 \frac{k_{ij}(x)}{k(x)} = \int_{\mathbb{R}^n} \left(a^2 \frac{f_{ij}(x-y)}{f(x-y)} + b^2 \frac{g_{ij}(y)}{g(y)} + 2ab \frac{f_i(x-y)}{f(x-y)} \frac{g_j(y)}{g(y)} \right) d\mu_x(y),$$

we have

$$(a + b)^2 \sum_{i,j=1}^n \alpha_i \alpha_j \frac{k_{ij}(x)}{k(x)} = \int_{\mathbb{R}^n} \sum_{i,j=1}^n \alpha_i \alpha_j \left(a^2 \frac{f_{ij}(x-y)}{f(x-y)} + b^2 \frac{g_{ij}(y)}{g(y)} + 2ab \frac{f_i(x-y)}{f(x-y)} \frac{g_j(y)}{g(y)} \right) d\mu_x(y).$$

Finally, for any given vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, and positive constants a, b we obtain the inequality

$$\begin{aligned} (a+b)^2 \sum_{i,j=1}^n \alpha_i \alpha_j \left(\frac{k_i(x)}{k(x)} \frac{k_j(x)}{k(x)} - \frac{k_{ij}(x)}{k(x)} \right) \leq \\ \sum_{i,j=1}^n \alpha_i \alpha_j \int_{\mathbb{R}^n} \left[a^2 \left(\frac{f_i(x-y)}{f(x-y)} \frac{f_j(x-y)}{f(x-y)} - \frac{f_{ij}(x-y)}{f(x-y)} \right) + \right. \\ \left. b^2 \left(\frac{g_i(y)}{g(y)} \frac{g_j(y)}{g(y)} - \frac{g_{ij}(y)}{g(y)} \right) \right] d\mu_x(y). \end{aligned} \quad (43)$$

Inequality (43) has important consequences. Indeed, the function $k(x)$, convolution of the two log-concave functions is log-concave. Consequently, the Hessian matrix \mathcal{M} of $-(a+b)^2 \log k(x)$ is a symmetric positive semi-definite matrix, which implies that, for any given vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$

$$- \sum_{i,j=1}^n \alpha_i \alpha_j (a+b)^2 (\log k(x))_{ij} = (a+b)^2 \sum_{i,j=1}^n \alpha_i \alpha_j \left(\frac{k_i(x)}{k(x)} \frac{k_j(x)}{k(x)} - \frac{k_{ij}(x)}{k(x)} \right) \geq 0.$$

Likewise, since the functions f and g are log-concave, the matrix \mathcal{N} with elements

$$n_{ij} = \int_{\mathbb{R}^n} \left[a^2 \left(\frac{f_i}{f} \frac{f_j}{f} - \frac{f_{ij}}{f} \right) (x-y) + b^2 \left(\frac{g_i}{g} \frac{g_j}{g} - \frac{g_{ij}}{g} \right) (y) \right] d\mu_x(y)$$

is a symmetric positive semi-definite matrix. Thus, inequality (43) asserts that the matrix $\mathcal{N} - \mathcal{M}$ is itself a symmetric positive semi-definite matrix. Consequently, the product matrix $(\mathcal{N} - \mathcal{M})(\mathcal{N} + \mathcal{M})$ is positive semi-definite, which implies that its trace is non-negative. Now, in view of classical properties of the trace of a matrix we obtain

$$\text{tr}(\mathcal{N} - \mathcal{M})(\mathcal{N} + \mathcal{M}) = \text{tr} \mathcal{N} \mathcal{N} + \text{tr} \mathcal{N} \mathcal{M} - \text{tr} \mathcal{M} \mathcal{N} - \text{tr} \mathcal{M} \mathcal{M} =$$

$$\text{tr} \mathcal{N} \mathcal{N} - \text{tr} \mathcal{M} \mathcal{M} = \sum_{i,j=1}^n n_{ij}^2 - \sum_{i,j=1}^n m_{ij}^2 \geq 0.$$

Finally, inequality (43) implies

$$\begin{aligned} (a+b)^2 \sum_{i,j=1}^n \left(\frac{k_i(x)}{k(x)} \frac{k_j(x)}{k(x)} - \frac{k_{ij}(x)}{k(x)} \right)^2 \leq \\ \sum_{i,j=1}^n \left\{ \int_{\mathbb{R}^n} \left[a^2 \left(\frac{f_i}{f} \frac{f_j}{f} - \frac{f_{ij}}{f} \right) (x-y) + b^2 \left(\frac{g_i}{g} \frac{g_j}{g} - \frac{g_{ij}}{g} \right) (y) \right] d\mu_x(y) \right\}^2. \end{aligned} \quad (44)$$

Applying once more Jensen's inequality to the squares into the right-hand side of inequality (44), and then proceeding as in the proof of Theorem 1 we easily arrive to inequality (6). Then, the cases of equality are found by the same argument of the one-dimensional proof. We conclude with the following

Theorem 6. *Let $f(x)$ and $g(x)$ be log-concave probability density functions with values in \mathbb{R}^n , with $n > 1$, such that both $J(f)$ and $J(g)$, as given by (1) are bounded. Then, $J(f * g)$ is bounded, and for any pair of positive constants a, b*

$$J(f * g) \leq \frac{a^4}{(a+b)^4} J(f) + \frac{b^4}{(a+b)^4} J(g) + 2 \frac{a^2 b^2}{(a+b)^4} H(f, g), \quad (45)$$

where

$$H(f, g) = \sum_{i,j=1}^n \int_{\{f>0\}} \frac{\partial_i f \partial_j f}{f} dx \int_{\{g>0\}} \frac{\partial_i g \partial_j g}{g} dx.$$

Moreover, there is equality in (45) if and only if, up to translation and dilation f and g are Gaussian densities, $f(x) = M_a(x)$ and $g(x) = M_b(x)$.

As for the one-dimensional case, given the random vector X distributed with density $f(x)$, $x \in \mathbb{R}^n$, consider the generic element of the second-order score function $\mathcal{H}(X)$, given by (24)

$$\Psi_{ij}(X) = \frac{f_i(X) f_j(X)}{f(X)} - \frac{f_{ij}(X)}{f(X)}.$$

Then, for each pair of i, j it holds the identity

$$\langle \Psi_{ij}(X) \rangle = \int_{\{f>0\}} \frac{\partial_i f \partial_j f}{f} dx, \quad \langle \Psi_{ij}(X)^2 \rangle = \int_{\{f>0\}} \left[\frac{\partial_i f}{f} - \frac{\partial_i f \partial_j f}{f^2} \right]^2 f dx.$$

Then, the standard inequality $\langle \Psi_{ij}(X)^2 \rangle \geq \langle \Psi_{ij}(X) \rangle^2$ gives

$$J(X) = \sum_{i,j=1}^n \langle \Psi_{ij}(X)^2 \rangle \geq \sum_{i,j=1}^n \langle \Psi_{ij}(X) \rangle^2 = \sum_{i,j=1}^n \left[\int_{\{f>0\}} \frac{\partial_i f \partial_j f}{f} dx \right]^2 \quad (46)$$

Using the Cauchy-Schwarz inequality, (46) gives

$$\begin{aligned} H(f, g) &= \sum_{i,j=1}^n \int_{\{f>0\}} \frac{\partial_i f \partial_j f}{f} dx \int_{\{g>0\}} \frac{\partial_i g \partial_j g}{g} dx \leq \\ &\left\{ \sum_{i,j=1}^n \left[\int_{\{f>0\}} \frac{\partial_i f \partial_j f}{f} dx \right]^2 \right\}^{1/2} \left\{ \sum_{i,j=1}^n \left[\int_{\{g>0\}} \frac{\partial_i g \partial_j g}{g} dx \right]^2 \right\}^{1/2} \leq \sqrt{J(f)} \sqrt{J(g)}. \end{aligned} \quad (47)$$

Hence, we can proceed as in the proof of Corollary 4 to obtain

Corollary 7. *Let X and Y be independent multi-dimensional random variables with log-concave probability density functions with values in \mathbb{R}^n , such that both $J(X)$ and $J(Y)$, as given by (1) are bounded. Then*

$$\frac{1}{\sqrt{J(X+Y)}} \geq \frac{1}{\sqrt{J(X)}} + \frac{1}{\sqrt{J(Y)}}.$$

Moreover, there is equality if and only if, up to translation and dilation X and Y are Gaussian densities.

3 A strengthened entropy power inequality

In this section, we will study the evolution in time of the functional $\Lambda(t)$ defined in (12), that is

$$\Lambda(t) = \Phi(f(t), g(t)) = H(f(t) * g(t)) - \kappa H(f(t)) - (1 - \kappa)H(g(t)).$$

Here, κ is a positive constant, with $0 < \kappa < 1$, while $f(x, t)$ (respectively $g(x, t)$) are the solutions to the heat equation (3) with diffusion constant κ (respectively $1 - \kappa$), corresponding to the initial data f and g , log-concave probability densities in \mathbb{R}^n . It is a simple exercise to verify that $\Lambda(t)$ is dilation invariant. This property allows to identify the limit, as $t \rightarrow \infty$ of the functional $\Lambda(t)$. For large times, the solution to the heat equation approaches the fundamental solution. This large-time behaviour can be better specified by saying that the solution to the heat equation (3) satisfies a property which can be defined as the *central limit property*. Suppose the initial density f in equation (3) is such that $(1 + |x|^\delta)f(x) \in L^1(\mathbb{R}^n)$ for some constant $\delta > 0$ (typically $\delta = 2$). Then, if

$$U(x, t) = \left(\sqrt{1 + 2t}\right)^n u(x \sqrt{1 + 2t}, t). \quad (48)$$

$U(x, t)$ tends in $L^1(\mathbb{R}^n)$ towards a limit function as time goes to infinity, and this limit function is a Gaussian function

$$\lim_{t \rightarrow \infty} U(x, t) = M_\kappa(x) \int_{\mathbb{R}^n} f(x) dx = M_\kappa(x). \quad (49)$$

This convergence property, as well as convergence in other stronger norms, can be achieved easily by resorting to Fourier transform, or by exploiting the relationship between the heat equation and the Fokker–Planck equation [7, 14] (cf. also [4] for recent results and references). We note that the passage $u(x, t) \rightarrow U(x, t)$ defined by (48) is dilation invariant, so that

$$\int_{\mathbb{R}^n} U(x, t) dx = \int_{\mathbb{R}^n} u(x, t) dx. \quad (50)$$

Coupling the dilation invariance of $\Lambda(t)$ with the central limit property, and remarking that $M_\kappa * M_{1-\kappa} = M_1$, gives

$$\begin{aligned} \lim_{t \rightarrow \infty} \Lambda(t) &= H(M_1) - \kappa H(M_\kappa) - (1 - \kappa) H(M_{1-\kappa}) = \\ &= -\frac{n}{2} [\kappa \log \kappa + (1 - \kappa) \log(1 - \kappa)]. \end{aligned} \quad (51)$$

If we differentiate $\Lambda(t)$ with respect to time, by the Bruijn's identity we obtain

$$\Lambda'(t) = I(f(t) * g(t)) - \kappa^2 I(f(t)) - (1 - \kappa)^2 I(g(t)). \quad (52)$$

In view of the inequality [5, 25]

$$I(f * g) \leq \frac{a^2}{(a+b)^2} I(f) + \frac{b^2}{(a+b)^2} I(g), \quad a, b > 0$$

with equality if and only if f and g are Gaussian densities, $\Lambda'(t) \leq 0$. Hence $\Lambda(t)$ is monotonically decreasing in time from $\Lambda(0)$ to $\Lambda(\infty)$, given by (51). Moreover, optimizing over κ , one shows that the non-negative quantity

$$H(f * g) - \kappa H(f) - (1 - \kappa) H(g) - \frac{n}{2} [\kappa \log \kappa + (1 - \kappa) \log(1 - \kappa)]$$

attains the maximum when

$$\kappa = \bar{\kappa} = \frac{\exp \left\{ \frac{2}{n} H(f) \right\}}{\exp \left\{ \frac{2}{n} H(f) \right\} + \exp \left\{ \frac{2}{n} H(g) \right\}}, \quad (53)$$

Evaluating $\Lambda(0) - \Lambda(\infty)$ in $\kappa = \bar{\kappa}$ gives the entropy power inequality (9). Note that this result holds for all probability densities.

Differentiating again with respect to time, from (52) we obtain

$$\Lambda''(t) = -\frac{1}{2} (J(f(t) * g(t)) - \kappa^3 J(f(t)) - (1 - \kappa)^3 J(g(t))). \quad (54)$$

Therefore, by inequality (37), if f and g are log-concave, $\Lambda''(t) \geq 0$, and the convexity property of $\Lambda(t)$ follows.

On the other hand, proceeding as in the proof of Corollary 4, we obtain from inequality (45) the bound

$$J(f(t) * g(t)) \leq \kappa^3 J(f(t)) + (1 - \kappa)^3 J(g(t)) - 2\kappa^2(1 - \kappa)^2 P(f(t), g(t)), \quad (55)$$

where

$$P(f, g) = \sqrt{J(f)} \sqrt{J(g)} - H(f, g) \geq 0 \quad (56)$$

in view of inequality (47). In addition, equality to zero holds if and only if both f and g are Gaussian densities.

Integrating (55) from t to ∞ , we obtain for the Fisher information of two log-concave densities the strengthened inequality

$$I(f(t) * g(t)) \leq \kappa^2 I(f(t)) - (1 - \kappa)^2 I(g(t)) - \kappa^2 (1 - \kappa)^2 \int_t^\infty P(f(s), g(s)) ds. \quad (57)$$

In fact, by the central limit property,

$$\lim_{t \rightarrow \infty} [I(f(t) * g(t)) - \kappa^2 I(f(t)) - (1 - \kappa)^2 I(g(t))] =$$

$$I(M_1) - \kappa^2 I(M_\kappa) - (1 - \kappa)^2 I(M_{1-\kappa}) = 0$$

Last, integrating (57) from 0 to ∞ we obtain for Shannon's entropy of the two log-concave densities the strengthened inequality

$$H(f * g) - \kappa H(f) - (1 - \kappa) H(g) - \frac{n}{2} [\kappa \log \kappa + (1 - \kappa) \log(1 - \kappa)] \geq \mathcal{P}_\kappa(f, g), \quad (58)$$

where

$$\mathcal{P}_\kappa(f, g) = \kappa^2 (1 - \kappa)^2 \int_0^\infty d\tau \int_\tau^\infty P(f(s), g(s)) ds \geq 0. \quad (59)$$

Choosing now $\kappa = \bar{\kappa}$ as given by (53) we end up with inequality (14), where

$$R(X, Y) = \exp \left\{ \frac{2}{n} \mathcal{P}_{\bar{\kappa}}(f, g) \right\} > 1. \quad (60)$$

Note that the term $R(X, Y)$ is related to the second-order score of the random vectors X and Y . Consequently, $R(X, Y) = 1$ if and only if both X and Y are Gaussian random vectors.

Remark 8. In general, the expression of the term $R(X, Y)$ is very complicated, due to the fact that it is given in terms of integrals of nonlinear functionals evaluated along solutions to the heat equations which depart from the densities of X and Y . It would be certainly interesting to be able to express the term $R(X, Y)$ (or to bound it from below) in terms of some distance of X and Y from the space of Gaussian vectors. This problem is clearly easier in one dimension, where one can use the remainder as given by inequality (42), namely as the sum of the two contributions of the type $J(f) - I^2(f)$. In this case, one would know if, for some distance $d(f, g)$ between two probability densities f and g and some positive constant C

$$J(f) - I^2(f) \geq C \inf_{M \in \mathcal{M}} d(f, M),$$

where \mathcal{M} denotes the space of Gaussian densities.

4 Conclusions

In this paper, we analyzed various inequalities for convolutions for log-concave densities. The main discovery is that log-concave densities satisfy a new inequality for convolutions which appears as the natural generalization of Shannon's entropy power (9) and Blachman–Stam (11) inequalities. This inequality is sharp, and it is the starting point for deriving Shannon's entropy power and Blachman–Stam inequalities in a strengthened form. It results in a clear way from the present analysis, that the behavior of the log-concave density functions with respect to convolutions deserves further investigations.

Acknowledgment: This work has been written within the activities of the National Group of Mathematical Physics of INDAM (National Institute of High Mathematics). The support of the project “Optimal mass transportation, geometrical and functional inequalities with applications”, financed by the Minister of University and Research, is kindly acknowledged.

References

- [1] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations, *Comm. Partial Differential Equations*, **26** 43–100 (2001).
- [2] K.I. Babenko, *An inequality in the theory of Fourier analysis*. Izv. Akad. Nauk SSSR, Ser. Mat **25** (1961), 531–542; English transl. Amer. Math. Soc. Transl. (2) **44** (1965), 115–128.
- [3] D. Bakry and M. Emery, Diffusions hypercontractives, *Sém. Proba. XIX*, Lect. Notes in Math. n. 1123, Springer Verlag, Berlin, 177–206 (1985).
- [4] J-P. Bartier, A. Blanchet, J. Dolbeault and M. Escobedo, *Improved intermediate asymptotics for the heat equation*. Appl. Math. Letters **24** (2011), 76–81.
- [5] N.M. Blachman, The convolution inequality for entropy powers. *IEEE Trans. Inform. Theory* **2**, 267–271, (1965).
- [6] H.J. Brascamp and E.H. Lieb, *Best constants in Young's inequality, its converse and its generalization to more than three functions*. Adx. Math. **20** (1976), 151–173.
- [7] J.A. Carrillo, G. Toscani, *Exponential convergence toward equilibrium for homogeneous Fokker-Planck-type equations*. Math. Methods Appl. Sci., **21** (1998), 1269–1286.

- [8] M.H.M. Costa, A new entropy power inequality, *IEEE Trans. Inf. Theory*, **IT-31**, (6) 751–760, (1985).
- [9] T.M. Cover, and J.A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.
- [10] D.R. Cox, H and D.V. Hinkley, *Theoretical Statistics*, Chapman & Hall, 1974
- [11] S. Dharmadhikari, and K. Joag-Dev, *Unimodality, Convexity, and Applications*, Academic Press, Boston (1988).
- [12] A. Dembo, A simple proof of the concavity of the entropy power with respect to the variance of additive normal noise, *IEEE Trans. Inform. Theory* **35**, 887–888, (1989).
- [13] A. Dembo, T.M. Cover, and J.A. Thomas, Information theoretic inequalities, *IEEE Trans. Inf. Theory*, **37**, (6), 1501–1518, (1991).
- [14] T. Goudon, S. Junca, and G. Toscani, *Fourier-based distances and Berry-Esseen like inequalities for smooth densities*. *Monatsh. Math.* **135** (2002) 115–136.
- [15] D. Guo, S. Shamai, and S. Verdú, Mutual information and minimum mean-square error in Gaussian channels, *IEEE Trans. Inf. Theory*, **51**, (4) 1261–1282, (2005).
- [16] D. Guo, S. Shamai, and S. Verdú, A Simple Proof of the Entropy-Power Inequality *IEEE Trans. Inf. Theory*, **52**, (5) 2165–2166, (2006).
- [17] P.L. Lions, G. Toscani, *A strengthened central limit theorem for smooth densities*. *J. Funct. Anal.* **129** (1995), 148–167.
- [18] M. Madiman, and A. Barron, Generalized entropy power inequalities and monotonicity properties of information, *IEEE Trans. Inf. Theory*, **53**, (4) 2317–2329, (2007).
- [19] H.P. McKean Jr., Speed of approach to equilibrium for Kac’s caricature of a Maxwellian gas, *Arch. Rat. Mech. Anal.* **21**, 343–367, (1966).
- [20] A.W. Marshall, and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, Orlando (1979).
- [21] A. Prékopa, *Logarithmic concave measures with application to stochastic programming*. *Acta. Sci. Math. Szeged* **32** (1971), 301–315.

- [22] O. Rioul, A simple proof of the entropy-power inequality via properties of mutual information, in *Proceedings of the 2007 IEEE International Symposium on Information Theory (ISIT 2007)*, Nice, France, June 24-29-th, 2007.
- [23] G. Savaré, and G. Toscani, The concavity of Renyi entropy power, *IEEE Transactions on Information Theory*, **60** (5) (2014), 2687–2693
- [24] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27**, 379–423, 623–656, Jul./Oct. (1948).
- [25] A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inf. Contr.* **2**, 101–112, (1959).
- [26] G. Toscani, Sur l’inégalité logarithmique de Sobolev, *CRAS* **324**, Série I 689–694 (1997).
- [27] G. Toscani, An information-theoretic proof of Nash’s inequality. *Rend. Lincei Mat. Appl.*, **24** (2013), 83–93.
- [28] G. Toscani, Lyapunov functionals for the heat equation and sharp inequalities. *Atti Acc. Peloritana Pericolanti, Classe Sc. Fis. Mat. e Nat.*, **91** (2013), 1–10.
- [29] G. Toscani, Heat equation and convolution inequalities, *Milan J. Math.* (in press) (2014)
- [30] C. Villani A short proof of the concavity of entropy power, *IEEE Trans. Info. Theory* **46**, (4) 1695-1696, (2000).
- [31] R. Zamir and M. Feder, A generalization of the entropy power inequality with applications, *IEEE Trans. Inf. Theory*, **39**, (5) 1723–1728, (1993).